

University of Wollongong

Research Online

Faculty of Engineering and Information
Sciences - Papers: Part A

Faculty of Engineering and Information
Sciences

2011

Learning-based prostate localization for image guided radiation therapy

Luping Zhou

University of Wollongong, lupingz@uow.edu.au

Shu Liao

University of North Carolina

Wei Li

University of North Carolina

Dinggang Shen

University of North Carolina at Chapel Hill

Follow this and additional works at: <https://ro.uow.edu.au/eispapers>



Part of the [Engineering Commons](#), and the [Science and Technology Studies Commons](#)

Recommended Citation

Zhou, Luping; Liao, Shu; Li, Wei; and Shen, Dinggang, "Learning-based prostate localization for image guided radiation therapy" (2011). *Faculty of Engineering and Information Sciences - Papers: Part A*. 1775. <https://ro.uow.edu.au/eispapers/1775>

Research Online is the open access institutional repository for the University of Wollongong. For further information contact the UOW Library: research-pubs@uow.edu.au

Learning-based prostate localization for image guided radiation therapy

Abstract

Accurate prostate localization is the key to the success of radiotherapy. It remains a difficult problem for CT images due to the low image contrast, the prostate motion, and the uncertain presence of rectum gas. In this paper, a learning based framework is proposed to improve the accuracy of prostate detection in CT. It adaptively determines distinctive feature types at distinctive image regions, thus filtering out features that are salient in image appearance, but irrelevant to prostate localization. Furthermore, an image similarity function is learned to make the image appearance distance consistent with the underlying prostate alignment. The efficacy of our proposed method has been demonstrated by the experiment.

Keywords

prostate, image, learning, localization, therapy, radiation, guided

Disciplines

Engineering | Science and Technology Studies

Publication Details

Zhou, L., Liao, S., Li, W. & Shen, D. (2011). Learning-based prostate localization for image guided radiation therapy. 8th IEEE International Symposium on Biomedical Imaging: From Nano to Macro (pp. 2103-2106). Chicago, United States:

LEARNING-BASED PROSTATE LOCALIZATION FOR IMAGE GUIDED RADIATION THERAPY

Luping Zhou, Shu Liao, Wei Li, Dinggang Shen

IDEA lab, Department of Radiology and BRIC, University of North Carolina at Chapel Hill, U.S.A

ABSTRACT

Accurate prostate localization is the key to the success of radiotherapy. It remains a difficult problem for CT images due to the low image contrast, the prostate motion, and the uncertain presence of rectum gas. In this paper, a learning based framework is proposed to improve the accuracy of prostate detection in CT. It adaptively determines distinctive feature types at distinctive image regions, thus filtering out features that are salient in image appearance, but irrelevant to prostate localization. Furthermore, an image similarity function is learned to make the image appearance distance consistent with the underlying prostate alignment. The efficacy of our proposed method has been demonstrated by the experiment.

Index Terms— Prostate Localization, Feature Selection, Image Similarity Learning, IGRT

1. INTRODUCTION

Prostate cancer is the second leading cause of male cancer death in US. As a non-invasive approach, external beam radiation therapy has been shown to provide an effective treatment for patients. The success of radiotherapy depends on the accurate localization of the target organ, because the prostate tumor is surrounded by healthy tissues that can also be harmed by radiation. In clinic, it is necessary to minimize the dosage applied to the bladder and rectum to mitigate post-treatment complications. For this purpose, the treatment is spread out over a weeks-long series of daily fractions and involves image-guided radiation therapy (IGRT) for the adjustment of patient setup. The whole process comprises two stages. In the planning stage, the prostate is manually delineated on the planning CT image. After that, in the daily treatment stage, a new CT image is acquired prior to each treatment. The prostate in the new treatment image is localized for the determination of the patient setup, so that the treatment beams can be adjusted accordingly to control the dosage in the patient space.

However, the localization of the prostate in CT images is challenging because of the following reasons. Firstly, the contrast of the prostate and the surrounding organs are not sharp, which introduces ambiguities of the boundaries of the prostate. Secondly, due to the motion of the prostate across different treatment images, there are heavy variations in the spatial relationship between the prostate and its surrounding soft tissues. This gives difficulties to prostate segmentation methods using profile-based features [1]. Region-based approaches using probability distribution functions are often used to handle this problem [2, 3]. However, these approaches are also criticized for their segmentation precision due to the ignorance of local image patterns [4]. Thirdly, the uncertain presence of rectum gas makes the appearance of the treatment images inconsistent. This may cause failure of some common segmentation and registration methods. Therefore, the rectum gas is often separately treated in the existing approaches. For example, Davis et al., proposed a “deflation approach” [5] to remove the inconsistency of gas across treatment images for deformable registration of the prostate.

In this paper, we propose a learning based framework for accurate localization of the prostate in CT treatment images. The images are modeled by a pool of local features. The proposed method can flexibly identify the best feature types at distinctive regions in the images for prostate localization. The selected features at selected locations are utilized to train an image similarity function that links the image appearance distance with the degree of the prostate alignment in the planning and treatment images. Therefore, when a new treatment image is given, its optimal alignment with the planning image can be effectively sought via the learned image similarity function. One advantage of our proposed learning approach is to be able to adaptively filter out image features that are salient in image appearance yet having little to do with the localization of the prostate, for example, the rectum gas.

In this paper, we focus on the estimation of translation parameters of the prostate. The reason is based on our observations that for the same subject, the majority of the intra-treatment prostate shape difference is caused by the translation of the prostate. Our experiment also shows that the mere removal of translation has been able to achieve an overlap ratio above 75% between the planning and the treatment prostates. The small remaining shape difference is mostly

This is an invited paper in the special session on “Imaging the Prostate: Histology and In vivo”. This work was supported in part by NIH grants R01 CA140413 and R01 EB008374. The image data and expert contours used in this study were provided by the Department of Radiation Oncology, UNC-Chapel Hill under support of NIH grants R01 RR018615 and R44/43 CA119571, E.L. Chaney PI.

caused by rotations. Only a very small amount of local deformations are involved during the treatment of the prostate. Please note that although our proposed method is demonstrated by estimating the translation parameters, it can be directly applied to the estimation of rotation parameters, and easily generalized to the estimation of the local deformation parameters.

2. METHOD

2.1. Image preprocessing

To remove the whole-body rigid motion of the patient, the bones (more specifically, pelvis, which remains fixed) in the treatment images are first rigidly aligned with that in the planning image. After the well alignment of bones, the intra-treatment prostate shape difference is likely to be caused by the motion of the prostate itself. In the following, we show how to accurately estimate the parameters that characterize this motion. Our method can be factorized into five steps: feature extraction, enrichment of training samples, distinctive feature selection, learning of image similarity, and prediction on the new treatment image, as introduced one-by-one below.

2.2. Feature extraction

To capture the characteristic of the local image appearance, at each voxel, 14 Haar features, 9 histogram of gradient (HOG) features, and 30 local binary pattern features are extracted from its 20×20 rectangle neighborhood. These features describe the intensity statistic (Haar features), intensity gradients (HOG), and texture (local binary patterns [6]) of the local image appearance, respectively. The Haar features are constructed in multi-resolution to suppress the noise at different levels. The HOG features are the 3×3 histogram of the orientations of the intensity gradients. Eventually a CT image is represented by 53 feature maps of the same image size, each of which belongs to a specific type of the above features. Please note that although 2D image features are used for each slice, our proposed method works on 3D images for translation parameter estimation.

2.3. Enrichment of the training samples

Due to the rich features at each image voxel, the number of the training images (the treatment images plus the planning image) needs to be large enough to accommodate a robust learning. However, on the other hand, in order to increase the clinic viability, the number of treatment images used for training should be minimized. Therefore, a strategy of perturbation is used to increase the number of training images as well as keeping a small number of the required treatment images. In particular, the bone-aligned treatment images (for training) are further brought into the space of the planning image by a translation that maximizes the prostate overlaps in the planning and the treatment images. After that, in the space of the planning image, each translated treatment image is used as a seed to generate a series of synthetic images by small random translation perturbations. The process is illustrated in

Fig 1. Clearly, the seed treatment image is best aligned with the planning image, compared with all of its perturbation-derived synthetic images. The feature maps of the synthetic images are simply obtained by translating the feature maps of the corresponding seed treatment images.

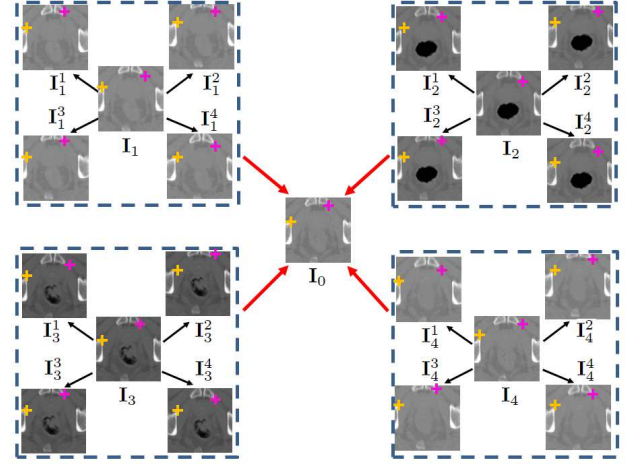


Fig. 1: The enrichment of the training data consists of two steps. First, each of the treatment image is aligned with the planning image I_0 by translation (the red arrows). Second, the position of the aligned treatment image I_t ($t = 1, \dots, 4$) is randomly perturbed to generate more training data (the black arrows). To show the translation clearly, two landmarks are indicated by the cross-hairs with different colors in all images consistently.

2.4. Feature selection

Even with the enrichment of the training samples, the number of the image features is still much larger than the number of the training images. To reduce the influence of “curse of dimensionality”, only a small subset of informative features have to be identified for the training. The distinctive features are expected to have the following property: when comparing the planning and the treatment images, the feature difference should be consistent with the overlap degree of the segmented prostates. That is, if the overlap of the segmented prostates is small, the feature difference should be large, and vice versa. The most straightforward measure of the prostate overlap is the Dice ratio (referred as overlap ratio in this paper), which is an implicit function of the translation parameters. However, this function may be complicated and have many local minima, as different parameters may lead to similar overlap ratios. To overcome this problem, we explicitly define a score function that has a simpler form of the translation parameters, as well as reflecting the overlap degree of the segmented prostates. The score function is elaborated as follows.

Let us denote the planning image as I_0 and the aligned treatment image as I_t ($t = 1, \dots, m$), where m is number of the treatment images used for training. Without loss of generality, assume that I_t^i ($i = 1, \dots, n$) is a training image generated from a seed treatment image I_t by perturbation de-

scribed in Section 2.3, where n is the number of perturbations for each seed image. The image appearances of \mathbf{I}_t^i and \mathbf{I}_t are different only due to the translation perturbation. Let us denote the translation parameters of \mathbf{I}_t^i from its seed image \mathbf{I}_t as $\Delta\theta_t^i$. The score function of the parameters $\Delta\theta_t^i$ is defined as

$$\mathcal{S}(\Delta\theta_t^i) = \gamma_t e^{-\frac{1}{2}(\Delta\theta_t^i)^\top \Sigma^{-1} \Delta\theta_t^i}, \quad (1)$$

where γ_t is the overlap ratio of the prostates in the aligned treatment image \mathbf{I}_t ($t > 0$) and the planning image \mathbf{I}_0 . The parameter Σ is a 3×3 diagonal matrix, whose diagonal elements are the variances of translation parameters learned from the training images.

The score function is composed of two parts: an exponential function of $\Delta\theta_t^i$, and a modulate coefficient γ_t . To understand the score function, let us consider the exponential function first. When the position of the training image \mathbf{I}_t^i coincides with the position of its seed treatment image \mathbf{I}_t , the response of the exponential function is 1.0, indicating the best alignment between \mathbf{I}_t^i and \mathbf{I}_0 . When \mathbf{I}_t^i deviates from \mathbf{I}_t by perturbing translations, the response of the exponential function decreases, indicating an increasing misalignment between \mathbf{I}_t^i and \mathbf{I}_0 . Note that for different treatment images \mathbf{I}_t ($t = 1, \dots, m$), the degree of their best alignments with the planning image may be quite different. This is because some treatment images may undergo bigger rotation and local deformations than the others. Such a difference is coped with by the parameter γ_t used to weight the exponential function. If \mathbf{I}_t aligns with \mathbf{I}_0 well, γ_t has a big value. Otherwise, γ_t has a small value. To facilitate the explanation, we call this score as Gauss score.

Two factors are catered in our feature selection: i) different image regions may contribute differently to the prostate localization; and ii) the best feature type in different image regions may be different. In other words, our feature selection is essentially selecting the best feature types at the distinctive locations. Please note that it is possible for more than one feature types to be selected at a distinctive location. Therefore, in our method, a “feature” refers to a pair (f, \mathbf{x}) , where f is a feature and \mathbf{x} is the location. The selection procedure is as follows. For a given feature (f, \mathbf{x}) , a feature distance vector is constructed as $\vec{\mathbf{F}}(f, \mathbf{x}) = [D_1^1(f, \mathbf{x}), \dots, D_t^i(f, \mathbf{x}), \dots, D_m^n(f, \mathbf{x})]$, where $D_t^i(f, \mathbf{x}) = f(\mathbf{I}_t^i(\mathbf{x})) - f(\mathbf{I}_0(\mathbf{x}))$. The feature distance vector indicates the distances between every training image \mathbf{I}_t^i and the planning image \mathbf{I}_0 measured by the feature (f, \mathbf{x}) . Also, a score vector is constructed as $\vec{\mathbf{S}} = [\mathcal{S}(\Delta\theta_1^1), \dots, \mathcal{S}(\Delta\theta_t^i), \dots, \mathcal{S}(\Delta\theta_m^n)]$. A distinctive feature (f, \mathbf{x}) should have highly correlated $\vec{\mathbf{F}}(f, \mathbf{x})$ and $\vec{\mathbf{S}}$, and therefore can be identified by thresholding the correlation coefficients of $\vec{\mathbf{F}}(f, \mathbf{x})$ and $\vec{\mathbf{S}}$.

2.5. Image similarity learning

The above feature selection step selects individual features, each of which measures image difference at a local (voxel)

level. In the following, the image difference is further estimated at a global level by combining the selected local features nonlinearly. In this way, the interactions of individual features are also taken into consideration.

Similar to the criterion used for individual feature selection, the image appearance similarity should also reflect the degree of underlying prostate overlaps. For this purpose, an image similarity function \mathcal{M} is learned as follows. Let us denote the selected features as $(\tilde{f}_l, \tilde{\mathbf{x}}_l)$, where $l = 1, \dots, L$. For each training image \mathbf{I}_t^i , a vector $\vec{\mathbf{F}}_t^i = [D_t^i(\tilde{f}_1, \tilde{\mathbf{x}}_1), \dots, D_t^i(\tilde{f}_L, \tilde{\mathbf{x}}_L)]$ is constructed. The mapping \mathcal{M} is obtained by regression, taking $\vec{\mathbf{F}}_t^i$ as the multivariate input and the corresponding Gauss score $\mathcal{S}(\Delta\theta_t^i)$ as the scalar output: $\mathcal{M} : \vec{\mathbf{F}}_t^i \mapsto \mathcal{S}(\Delta\theta_t^i)$. In our method, the Support Vector Regression (SVR) with a Gaussian RBF kernel $k(\vec{\mathbf{F}}_{t_1}^{i_1}, \vec{\mathbf{F}}_{t_2}^{i_2}) = \exp(-\frac{\|\vec{\mathbf{F}}_{t_1}^{i_1} - \vec{\mathbf{F}}_{t_2}^{i_2}\|^2}{2\sigma^2})$ is employed. A grid-search in a predefined range is used to choose the optimal hyper-parameters used in the SVR. The scale of the hyper-parameter σ used in the RBF kernel is determined by the average distance between every two feature distance vectors. As the last step in the training stage, SVR links the estimation of Gauss score (degree of prostate overlap) with the image appearance features at selected locations.

2.6. Prediction

When a new treatment image comes, given a translation θ , the prostate overlap between the translated planning image and the new treatment image can be measured by the image similarity function $\mathcal{M}(\vec{\mathbf{F}})$. The translation θ that generates the biggest similarity $\mathcal{M}(\vec{\mathbf{F}})$ is taken as our estimated optimal translation θ^* . In our application, as the target is to localize the prostate in the new treatment image, the reversed translation $-\theta^*$ is applied to bring the segmented planning prostate to the space of the new treatment image. Consequently, the position of the prostate in the new treatment image is estimated by the translated planning prostate.

3. RESULTS

Our proposed method is tested on CT image series comprising different time points for several patients. The bones in the treatment images are rigidly aligned with the bones in the planning image. The proposed method is trained using the planning image and the first 5 treatment images, and tested on the rest treatment images. To enrich the training samples, 40 random perturbations are applied on each of the planning and the 5 training treatment images, which gives rise to totally 240 training samples. Each training sample is associated with a Gauss score to indicate the degree of its prostate alignment with the planning image. By thresholding the correlations between the feature distances and the Gauss scores, above 900 features are selected, most of which are Haar features located

on the fringe of the prostate opposite to the rectum. Fig. 2 shows some of the selected discriminative positions.

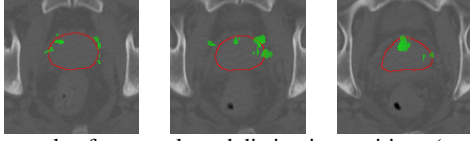


Fig. 2: Example of some selected distinctive positions (green points) by thresholding the correlation coefficients between the feature distances and the Gauss scores. The prostate in the planning image is delineated in red.

Table 1 shows some example results of the test treatment images. The second column exhibits the overlap ratios between the test treatment prostates and the planning prostate before the translations. The third column shows the best overlap ratios that can be achieved by translations only between the test treatment prostates and the planning prostate, which are used as ground truth. The best overlap ratios are obtained by optimizing the translations of the manually segmented prostates between the test and the planning images. The fourth column shows the overlap ratios between the test treatment prostates and the planning prostate after our estimated translations. It can be clearly seen that our proposed method can accurately estimate the translation parameters. The overlap ratios achieved after our estimated translations are very close to the expected overlap ratios. Their differences are only 0.0255 by average with a standard deviation of 0.0299. The efficacy of our proposed method is also visible in Fig. 3. Although the appearance differences of the test treatment images are large due to the uncertain presence of rectum gas, our method can always improve the prostate alignment from the original position. It can be observed that, after estimation, the planning prostate can be better aligned with the prostates in the treatment images.

Table 1: Results evaluated by overlap ratios

Image ID	Original Overlap	Expected Overlap	Estimated Overlap
1	0.4143	0.7860	0.6942
2	0.7019	0.7888	0.7800
3	0.2811	0.7697	0.7559
4	0.4020	0.8086	0.7802
5	0.5378	0.8815	0.8815
6	0.5203	0.8427	0.8082
7	0.4300	0.8274	0.8133
8	0.3682	0.7406	0.7373
9	0.2721	0.7117	0.7117
10	0.4180	0.7952	0.7346

4. CONCLUSION

In this paper, a learning based framework is proposed for accurate prostate localization in CT images. By a designed fea-

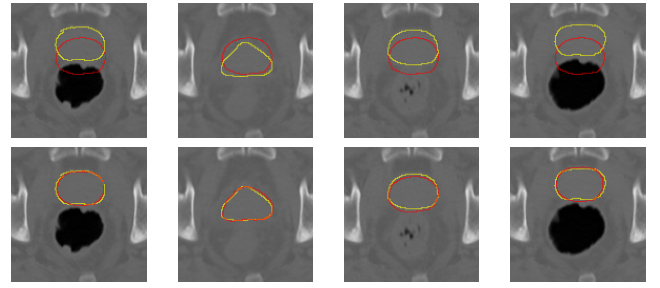


Fig. 3: Typical performance of our method in estimating the prostate locations in the treatment images. Row 1: The original treatment images, along with their respective manually-segmented prostates (yellow curves), as well as the planning prostate (red curve) before alignment. Row 2: The final alignment of the planning prostate (red curve) with each treatment image by our method; note that the yellow curves are the same as those in Row 1.

ture selection procedure, distinctive image regions with selected types of features are identified to train an image similarity function that effectively reflects the degree of prostate alignment. The proposed method will be extended to the estimation of local deformation parameters in our future work.

5. REFERENCES

- [1] F. A. Cosio, “Automatic initialization of an active shape model of the prostate,” *Medical Image Analysis*, vol. 12, pp. 469–483, 2008.
- [2] J. V. Stough, R. E. Broadhurst, S. M. Pizer, and E. L. Chaney, “Regional appearance in deformable model segmentation,” in *Information Processing in Medical Imaging*, 2007, pp. 532–543.
- [3] D. Freedman, R. J. Radke, T. Zhang, Y. Jeong, D. M. Lovelock, and G. T. Chen, “Model-based segmentation of medical imagery by matching distributions,” *IEEE Trans. Med. Imaging*, vol. 24, pp. 281–292, 2005.
- [4] Q. Feng, M. Foskey, W. Chen, and D. Shen, “Segmenting ct prostate images using population and patient-specific statistics for radiotherapy,” *Medical Physics*, vol. 37, pp. 4121–4132, 2010.
- [5] B. C. Davis, M. Foskey, J. Rosenman, L. Goyal, S. Chang, and S. Joshi, “Automatic segmentation of intra-treatment ct images for adaptive radiation therapy of the prostate,” in *Proceedings of MICCAI*, 2005, pp. 442–450.
- [6] T. Ojala, M. Pietikainen, and T. Maenpaa, “Multiresolution gray-scale and rotation invariant texture classification with local binary patterns,” *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 24, pp. 971–987, 2002.